J Bus Psychol (2014) 29:593–604
DOI 10.1007/s10869-014-9350-0

ORIGINAL PAPER

Storytelling in the Selection Interview? How Applicants Respond to Past Behavior Questions

Adrian Bangerter · Paloma Corvalan · Charlotte Cavin

Published online: 16 March 2014

Abstract
Purpose Increased use of past behavior questions makes it important to understand applicants' responses. Past behavior questions are designed to elicit stories from applicants. Four research questions were addressed: How do applicants respond to past behavior questions, in particular, how frequent are stories? When applicants produce stories, what narrative elements do they contain? Is story production related to applicants' characteristics? Do responses affect interview outcomes?
Design/Methodology/Approach Using a database of 62 real job interviews, the prevalence of five types of applicants' response to past behavior questions were analyzed: *story*, *pseudo-story*, *exemplification*, *value/opinion*, and *self-description*. We also coded the narrative content of stories, distinguishing between *situations*, *tasks/actions*, and *results*. We analyzed relations between applicant characteristics (gender, age, personality, self-reported communication and persuasion skills, general mental ability) and response type. We used hierarchical multiple regression to predict hiring recommendations from response type.
Findings Stories were only produced 23 % of the time. Stories featured more narrative elements related to situations than tasks, actions, or results. General mental ability and conscientiousness affected response types, and men produced more stories than women. There were differences in the storytelling rate according to the type of competency. Stories and pseudo-stories increased hiring recommendations, and self-descriptions decreased them.
Originality/Value Behavioral interviews may not be conducive to storytelling. Recruiters respond positively to narrative responses. More research is needed on storytelling in the selection interview, and recruiters and applicants might need training on how to encourage and tell accurate and representative stories.

Keywords Selection interview · Storytelling · Communication · Narrative · Behavioral questions

A. Bangerter (✉) · P. Corvalan · C. Cavin
Institute of Work and Organizational Psychology, University of Neuchâtel, Emile-Argand 11, 2000 Neuchâtel, Switzerland
e-mail: adrian.bangerter@unine.ch

Introduction

Personnel selection research has traditionally focused on establishing the predictive validity of selection methods (Schmidt and Hunter 1998). Recently, research has also focused on the parameters of applicants' responses in selection situations. Findings have evidenced the sometimes unintended ways in which applicants respond to selection methods. For example, in personality testing, the issue of faking has been widely investigated (Morgeson et al. 2007), and this research has led to the development of sophisticated methods for identifying fakers or attenuating or correcting for faking (e.g., conditional reasoning tests; James et al. 2005). Many such methods are now routinely implemented in commercially available testing procedures. One might argue that a complete picture of the properties of a selection method requires understanding of both its psychometric properties and how applicants respond to the method. These two aspects seem especially relevant for the large-scale commercial or technological implementation of a selection method. However, these aspects are not equally well understood for all selection methods.

The structured selection interview is a case in point. Much is known about its psychometric properties. A host

of studies have shown that structured interviews are more valid and reliable selection instruments than unstructured interviews (Huffcutt and Arthur 1994; McDaniel et al. 1994). However, comparatively little is known about the response parameters of applicants in structured interview situations. This is a striking fact, especially because the principles of structured interviewing are becoming implemented in practice. Despite evidence that practitioners resist structured interviewing (Lievens and De Paepe 2004; Van der Zee et al. 2002), some aspects of structured interviewing are spreading in practice (Roulin and Bangerter 2012). Methods for implementing structured interviewing procedures are available to practitioners in various languages (Jetter 2008; Roulin et al. 2012, Schuler 1992). The competency-based interview (Kessler 2006) is featured in many interview training programs. Thus, a *technology* of structured interviewing is emerging, with rating scales, interviewer training manuals, and mnemotechnic aids to support interviewers' practices (e.g., the widespread STAR or situation-task-action-result guide to interviewer prompting, Kessler 2006). In such a situation, it is important to improve the evidence base about how the structured interview works in practice.

There are many features of structured interviews that may contribute to enhancing validity, but a particularly important one is the type of questions applicants get asked. Structured interviews eschew more traditional questions in favor of so-called behavioral questions. There are two kinds of behavioral questions, past behavior questions and situational questions (Campion et al. 1997). Both question types involve asking applicants questions about their actions in work situations. Such questions are calibrated to competencies that have been previously identified as job-relevant. Past behavior questions ask applicants to recount past situations where they performed well. An example might be *Can you tell me about an occasion where you had to deal with an angry client*? Situational questions (sometimes used when applicants differ in their amount of job experience) ask applicants to imagine a fictitious job-related scenario and explain what they would do in such a situation (e.g., *Imagine you receive a phone call from a client who is angry about an error on their bill. What do you do in such a situation?*).

By their very nature, past behavior questions are designed to elicit a coherent account of what the applicant did in a particular situation. Producing such an account involves *telling a story* about what one did in such a situation, as expressed in some advice books on how to construct behavioral interviews: "Responses should tell a complete story with a beginning, a middle, and an end" (Larson 2001, p. 75). A competent response to such a question may involve describing actions as following from the situational constraints, or framing particularly impressive outcomes as having being caused by one's actions. Moreover, recruiters' prompts are designed to help applicants provide complementary information about that situation, or about the task, actions or results that transpired (Kessler 2006), in other words, to "keep the person moving through the sequence of the story" (Prabakar Kamath 2009, p. 92). Thus, answering past behavior questions requires a measure of storytelling skill on the part of the applicant for there to be some material for the recruiter to analyze (as we will see, it is also necessary for the applicant to construe the question as an invitation to tell a story). As explained by Ralston et al. (2003, p. 8), "although interviewing offers many opportunities for applicant storytelling, this skill is especially vital when interviewers ask behavioral description questions because such questions demand answers in the form of stories".

As behavioral interviews spread, storytelling skills may become increasingly important in the future. However, there is little scientific evidence that applicants are able to construe behavioral questions appropriately and to systematically produce meaningful, effective stories when invited to do so. It is important to understand the response parameters to past behavior questions in order to identify typical situations where responses may be less representative of applicants' true level of mastery of a particular competency. One such situation may be when applicants are not skilled at producing a good story. In such a case, their response may lead to an underestimation of their skills. Another situation may be when applicants are able to fake (Levashina and Campion 2006) by embellishing or exaggerating a story about their accomplishments, thereby leading recruiters to overestimate their skills. More generally, although structured interviewing implies rating the content of stories, there may be situations in which the storytelling performance itself affects (or interferes with) the content of the story. A better understanding of the properties of applicants' responses to behavioral questions could lead to several improvements in interview practice. For example, it may be possible to teach applicants to produce meaningful and representative stories as part of training programs (Maurer et al. 2008; Ralston et al. 2003). And interviewers may be better trained to help applicants produce such stories (e.g., by appropriate prompting).

Here we report results of an initial study of storytelling in structured interviews. We investigated how successful past behavior questions are in eliciting stories and whether storytelling is related to interview outcomes. We analyzed a database of real interviews to investigate four questions: How do applicants respond to interview questions explicitly soliciting a story? What narrative elements do stories contain? What individual characteristics are related to storytelling propensity? And how are applicants' responses related to interview outcomes?

Storytelling: Implications for the Selection Interview

To answer our research questions, it is important to understand the storytelling process and its effects in detail. We thus review relevant research on storytelling in conversation from sociolinguistics, conversation analysis, and the psychology of language. These related fields of study have led to a detailed understanding of the process of storytelling as a joint accomplishment and the issues conversational participants must solve together in the course of that accomplishment.

A story is a description of a sequence of events experienced by one or more characters. The characters react to situations and change them through their actions, thereby creating new events that need to be dealt with, and so on until the conclusion of the story (Bruner 1990). The events are narrated as being temporally or causally coherent, for example, one event is depicted as following another or as having been caused by the actions of someone. Research on how language is used in social interaction (Clark 1996; Labov and Waletzky 1967; Sacks et al. 1974) has explored many aspects of storytelling. Three aspects particularly relevant for personnel selection are (1) the collaborative nature of storytelling, (2) the potential effects of stories on the audience and (3) individual differences in narrative skill.

Like conversation more generally (Clark 1996; Grice 1975), storytelling is a collaborative process. In conversation, participants need to coordinate their efforts to create and maintain, moment by moment, a shared understanding of the activity they are accomplishing together. Thus, any kind of conversational activity needs to be collaboratively ratified by all participants. This is often accomplished implicitly, without impinging on the main topical activity, as when participants say *okay* to signal their readiness to end a telephone conversation or *uh-huh* to signal their understanding of what the speaker has said (Bangerter and Clark 2003; Bangerter et al. 2004; Schegloff and Sacks 1973). When switching from one conversational activity to the next, for example when changing a topic, participants in conversation need to collaboratively solve a problem of relevance (Sperber and Wilson 1986). That is, any new topic must be made relevant to the topic at hand, and participants often actively suggest relevance (Button and Casey 1984) via expressions like *speaking of which* or *that reminds me of*.

Engaging in an episode of storytelling is no exception (Goodwin 1984), and conversational participants use so-called pre-sequences in order to jointly commit to storytelling activity (Schegloff 2007). An example of a pre-sequence is when A asks B *you know what?* and B responds *no, what?* Pre-sequences serve to establish A and B's willingness to engage in an activity where A will produce an extended turn at talk, thus departing from the usual rules governing turn-taking in conversation (Sacks et al. 1974). Although it is the narrator who does the bulk of talking in telling a story, how listeners respond affects the course of the story. Distracted listeners produce less back-channel responses, and this lack of feedback interferes with the storyteller's performance (Bavelas et al. 2000). Speakers may also enlist the aid of listeners in searching for a word or a name during a story (Goodwin 1987). The collaborative nature of storytelling raises at least two important questions for story production in selection interviews. First, because storytelling constitutes a departure from normal turn-taking in conversation, stories need to be adequately introduced. Second, recruiters' expressions of interest, amusement, or surprise in reaction to applicants' stories may affect the course of the story itself, with ensuing implications for reliability (Campion et al. 1997).

Stories have particular effects on audiences. First, they indirectly convey impressions. A well-crafted story may exemplify an applicant's mastery of a competency without making a direct claim and thus avoid attributions of immodesty. Second, stories constrain attribution processes. Attributions are common aspects of applicants' talk in the selection interview (Silvester 1997), and stories may influence recruiters' attributions by proposing particular framings of events and actions (Edwards and Potter 1993). Third, stories are often vivid. Many feature reported speech (Holt 1996), where narrators switch from describing events to directly quoting other actors, sometimes augmenting these re-enactments with a range of gestural or postural mimicry (Sidnell 2006). These properties may make stories more engaging or more likely to be remembered by recruiters. Fourth, stories narrated in detail may be perceived as difficult to fake by recruiters, and thus as credible signals of an applicant's mastery of a concept (Bangerter et al. 2012).

There are likely individual differences in narrative skill, and these differences may affect storytelling performance of applicants. Narrative skill, or the ability to "produce and comprehend causally and temporally structured plots" (Kemper 1984, p. 99), develops over the course of childhood. Age may affect storytelling abilities in the selection interview, especially because young applicants may have a less consolidated life story (Habermas and Bluck 2000) or because older adults produce better-structured stories (Pratt and Robins 1991). More experienced applicants may also have a larger base of anecdotes that are worth narrating. Independent of actual work experience, individual storytelling experience may also vary. Many stories are narrated repeatedly, either for the same audience or a different one. In such retellings, the basic story structure is often reused (Norrick 2000). Applicants with relatively more interviewing experience may therefore be more fluent at

storytelling. Many of these individual differences in storytelling abilities may interact with the actual content of the story to determine interviewers' ratings.

This Study

Storytelling is widespread in organizational life, for example as a component of organizational culture (Boje 2008) or as a recurrent activity in communities of practice (Bangerter et al. 2011; Orr 1996). It is also touted by consultants as a key leadership skill (Denning 2004). As shown by the above review, storytelling might constitute an important dimension of social interaction in the selection interview. Extant research on social interaction in the selection interview focuses on phenomena like similarity between interviewers and applicants, nonverbal behavior, or impression management (IM) (Posthuma et al. 2002). Storytelling might constitute a means by which IM strategies are implemented. For example, in a field study, Stevens and Kristof (1995) found that storytelling was common (exhibited by 23 of 24 interviewees) and served to implement several IM tactics like enhancing performance or describing how particular obstacles were overcome. Also, that past behavior questions tend to elicit self-promotion (Ellis et al. 2002) suggests that stories might constitute effective ways to implement self-promotion. But the link between IM and storytelling remains unclear. More generally, the question remains how IM tactics get implemented in various types of response that may or may not take the form of stories.

Another important issue is how applicants respond to past behavior questions. Such questions constitute useful selection tools to the extent that applicants are able to respond appropriately. Past behavior questions formulated according to best practice recommendations (e.g., Larson 2001) may seem unproblematic. Indeed, according to a common-sense view of communication, if applicants are asked the question *Can you tell me about a situation (…)?*, they should respond by producing a story. However, successful communication involves correctly interpreting intentions (Clark 1996; Grice 1975), and questions are open to multiple interpretations. It is well known from survey research methodology, for example, that standardizing wording of interview questions does not automatically eliminate problems of meaning (Suchman and Jordan 1990). Survey interviews are particularly prone to misconstruals by respondents (Clark and Schober 1991; Conrad and Schober 2005). Just like survey interview questions, selection interview questions may also carry ambiguous interpretations. As suggested above, stories in everyday conversation need to be introduced appropriately, and progress in narrating a story is accompanied, moment-by-moment, by responses from the audience. The restricted social interaction typical of a structured selection interview may not be rich enough in cues that are conducive to storytelling, which thrives as an informal process. As such, storytelling might constitute a dispreferred response in the selection interview setting and applicants might be reluctant to launch into a vivid, detailed story because they might not be sure such a level of detail is relevant to the interview. Given these concerns, it is important to understand how applicants construe past behavior questions and what types of response such questions elicit.

In sum, then, there is a need for a research program into storytelling and other applicant response types in the selection interview, in particular for the case of behavioral questions. The present study constitutes an initial investigation in this direction. Because little is known about this phenomenon, we employed an inductive research strategy guided by four open-ended research questions. Research Question 1 is: *How do applicants respond to past behavior questions, in particular, how frequent are stories?* Although past behavior questions are designed to elicit stories, it is unclear whether applicants systematically respond in kind. To answer this research question, we developed an inductive coding system, based partly on prior research and partly on an initial examination of responses, to distinguish and classify various types of responses, including stories. Research Question 2 is: *When applicants produce stories, what narrative elements do they contain?* This research question seeks to understand the content of stories as a necessary prelude to understanding how content can potentially be affected by recruiters' actions, e.g., prompts or follow-up questions. To answer this question, we coded content in stories according to the categories of the STAR (situation-task-action-result) model, because this model is widespread as a guideline for recruiters. Research Question 3 is: *Is story production related to applicants' characteristics?* To answer this question, we examined links between the propensity to produce stories and applicant characteristics. Research Question 4 is: *How are applicants' responses related to interview outcomes?* To answer this question, we investigated whether different kinds of response predict hiring recommendations of recruiters.

We investigated these questions using a database of 62 videotaped real job interviews. The database had been gathered as part of another research project on automatic sensing of applicant nonverbal behavior (Frauendorfer et al. 2013a, b). Applicants responded to a job advertisement for a research assistant position. All applicants participated in a selection procedure involving self-reports of communication and persuasion skills, the Big Five dimensions of personality, a general mental ability (GMA) test, and a structured interview focusing on four competencies (communication,

persuasion, organization, and stress management). A panel of professional recruiters made hiring recommendations for each candidate based on their self-report data and videotapes of the interviews. This procedure enabled us to get access to a large set of standardized structured interviews featuring past behavior questions for a real job position. Our analysis focused on coding the types of responses applicants produced to these questions. For those responses involving storytelling, we coded the narrative elements of the stories produced. We also controlled for applicants' responsiveness, i.e., how many words they produced, because this may affect hiring recommendations (e.g., via attributions of verbal dominance, Schmid Mast 2002).

One advantage of this database is the fact that applicants respond to four behavioral questions on different competencies, which constitute a wide range of stimuli. Another advantage is the fact that the interviewer systematically refrained from prompting applicants. Thus, applicants' responses constitute a stringent baseline profile of a spontaneous response to a past behavior question. A disadvantage of the database is the position itself, which naturally restricted the population to younger university students, thus limiting variation on variables like age or experience.

Method

Participants

There were 62 applicants (45 men, 17 women, mean age 23.7 years, *SD* = 3.8, 59 were students), recruited for the study at a French-speaking Swiss university. They had on average already participated in 3.1 selection interviews (*SD* = 1.3).

Procedure

Applicants responded to a job advertisement for a research assistant position, the main activity of which was to recruit participants for another study. On arrival in the lab, they took a general mental ability test and filled out personality and skills questionnaires, before undergoing a behavioral interview (average duration 11 min). As part of the interview questions, all applicants were asked four past behavior questions about their competencies of communication, persuasion, organization, and stress management (these competencies were mentioned in the advertisement). The question wording (as recommended by Larson 2001) was *Can you tell me about a situation where you showed your competency for X*? Applicants answered the question in the way they saw fit. When the applicant was done, the interviewer did not ask any follow-up questions, but moved on to the next question. Interviews were videotaped and responses to past behavior questions were transcribed word-for-word (but excluding disfluent speech).

Measures

Applicant Communication Skills

Self-reported communication skills were assessed via 13 items based on the Social Skills Inventory (Riggio 1986). An example item is *In general I communicate in a clear manner*. Items were answered on a Likert scale from 1 (*not at all*) to 5 (*very much*), *alpha* = .89.

Applicant Persuasion Skills

Self-reported persuasion skills were assessed via 6 items based on the Social Skills Inventory (Riggio 1986). An example item is *I often succeed in selling my point of view*. Items were answered on a Likert scale from 1 (*not at all*) to 5 (*very much*), *alpha* = .79.

Big Five

Applicants responded to 60 items measuring the five traits of the NEO PI-R (Costa and McCrae 1992): neuroticism (alpha = .83), extraversion (alpha = .72), openness (alpha = .71 with one item removed), agreeableness (alpha = .67 with one item removed) and conscientiousness (alpha = .89). Items were formulated as assertions (e.g., *I try to be courteous to everyone I meet* for agreeableness) which participants endorsed on a Likert scale from 1 (*not at all*) to 5 (*very much*).

Applicant GMA

GMA was assessed via the Wonderlic Personnel Test (Wonderlic 2001), which measures vocabulary, arithmetic reasoning, and spatial ability. Applicants had 12 min to answer up to 50 questions. The test score is the percentage of questions correctly answered.

Hiring Recommendation

Hiring recommendations were made by five professional recruiters who had between 2 and 10 years of experience in recruiting. Recommendations were made on a scale of 0 % (weakest recommendation) to 100 % (strongest recommendation). One recruiter viewed and rated videotapes of all applicants. The others viewed and rated a subset of the videotapes, such that each applicant was rated by three recruiters; the mean of the three ratings was computed. Recruiters also had access to applicants' NEO-FFI scores, the scores of the self-reports of communication and

persuasion skills and GMA scores. Interrater reliability was computed via intraclass correlations, ICC[1] = .50, ICC[2] = .75, $F = 4.45$, $p < .05$.

Coding

Responsiveness

We measured how responsive applicants were by counting the mean number of words for each applicant's response to each past behavior question. We then aggregated over all four questions to produce a mean score per applicant.

Response Type

We viewed the videotapes and coded whether or not applicant's responses featured five types of discourse: *story*, *pseudo-story*, *exemplification*, *value/opinion*, and *self-presentation*. A story was defined as a set of events related to a unique past episode, characterized by a unity of time or action, which constituents often linked by temporal markers (e.g., *then*). An example is *We started with nothing and we organized an evening where we convinced owners of a night club […] we had to take care of all the details. And it worked perfectly, we even did it again 3, 4 months afterward*. A pseudo-story was defined as a description of a generic situation or recurrent set of similar situations, without unity of time or action. It differs from a story in that it is a description of several events rather than of a unique event. As a result, pseudostories are typically more abstract than stories. An example is *Other situations where people left without paying. […] Someone who left with 120 francs that was 120 francs less pay for us. It was stressful because we had about 5 min to watch the camera, write down the license plate, call the police […]*. Exemplification was defined as a part of a pseudostory featuring a brief mention of specific contextual information, often marked by *for example*. An example is *When I prepare my exams and when I do other jobs as a student. For example, I just finished a two-week job as cook's helper and you need to be rather thorough*. Stories, pseudostories, and exemplification constitute *narrative* response types. However, applicants may also respond via decontextualized assertions. We distinguished two such response types. One such response type involves assertions about the applicant's values or opinions. An example is *If you love someone you need to show it and you need to explain your feelings*. Another decontextualized assertion is a self-description of a personal attribute, for example *That's something for which I am talented*.

Response types were not mutually exclusive, e.g., an applicant might tell a story and then express an opinion, and so the presence or absence of each response type in a response was coded independently (except for exemplifications which were associated by definition with pseudostories). Interrater agreement, based on double-coding of 24 responses, was high (Cohen's kappa varied between .74 and 1). We computed the proportion of responses (out of 4) which featured each response type (i.e., scores could be either 0, .25, .50, .75, or 1).

Story Content

We segmented each story into constituent utterances (typically corresponding to a clause with a subject, verb, and object) and coded each utterance, distinguishing between descriptions of the *situation* (e.g., *We started with nothing*), the *task/actions* undertaken by the applicant (e.g., *I asked her if she could wait a minute*), *results* (e.g., *And it worked perfectly*), and *other* narrative content (e.g., *You need to be a bit of an artist*). Interrater agreement, based on double-coding of 49 utterances, was high (kappa = .86).

Results

Descriptive statistics of main study variables appear in Table 1. We present main results according to Research Questions 1–4.

To answer Research Question 1, we analyzed the frequency of different response types. The most frequent response type (Table 1) is pseudostory, followed by values/opinions, self-descriptions, stories, and exemplifications. Proportion of stories was strongly negatively correlated with proportion of pseudostories, suggesting that applicants tended to produce either stories or pseudostories. Proportion of stories, pseudostories, and values/opinions were correlated with responsiveness, suggesting that these response types required more words.

To answer Research Question 2, we analyzed the 57 stories that were produced. Utterances focused significantly more often on situations ($M = 6.8$, $SD = 4.7$) than on tasks/actions ($M = 3.2$, $SD = 3.2$), results ($M = 3.2$, $SD = 3.2$), or other ($M = 3.2$, $SD = 3.2$) narrative elements. A repeated-measures ANOVA with utterance type as a four-level within subjects variable revealed a significant main effect, $F(2.5, 149.6) = 36.1$, $p < .001$ (corrected for violation of sphericity by the Huynh–Feldt method, Field 2009). Simple contrasts (with situation as the reference category) indicated that all other utterance types were less frequent than situational descriptions.

To answer Research Question 3, we investigated relations between applicant characteristics and the proportion of stories. We also investigated relations between applicant characteristics and the proportion of other response types. Table 1 shows that there are no significant correlations

Table 1 Means (*M*), standard deviations (*SD*) and correlations for main study variables

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. Gender	1.27	.45	–																
2. Age	23.72	3.76	.10	–															
3. Int. exp.	3.06	1.32	−.11	.31*	–														
4. E	3.98	.49	−.15	−.16	.16	–													
5. O	3.69	.55	.10	.23	.40**	.07	–												
6. N	2.25	.61	−.26*	.14	.16	−.42**	.07	–											
7. A	4.11	.46	.06	−.09	−.06	.40**	.09	−.29*	–										
8. C	4.10	.64	.14	−.02	.07	.26*	−.04	−.50***	.32*	–									
9. Comm.	3.82	.59	.23	−.16	.22	.34**	.24	−.47***	.45***	.53***	–								
10. Persuas.	3.98	.55	.25*	−.04	.19	.26*	.25	−.28*	.16	.31*	.48**	–							
11. GMA	.69	.14	−.16	−.41**	−.10	−.09	.00	.11	−.02	−.18	.08	−.11	–						
12. Resp.	128.46	56.57	.43**	.24	.17	.11	.19	−.31*	.19	.23	.11	.10	−.18	–					
13. Stories	.23	.22	.22	.04	.07	.14	.07	−.15	−.00	.07	.01	.09	−.26	.46***	–				
14. Pseudo.	.76	.24	−.18	−.11	−.09	.05	−.01	−.01	.10	−.01	.04	−.03	.36**	.28*	−.72***	–			
15. Exempl.	.15	.21	−.09	−.20	−.18	−.11	−.04	.01	−.05	−.09	−.09	−.24	.33**	.04	−.26*	.29*	–		
16. V/O	.29	.27	−.03	.40**	.21	.18	−.01	−.13	.24	.27*	.12	.11	−.22	.37**	.00	−.05	−.11	–	
17. Self-d.	.29	.31	.14	.05	.00	−.15	−.07	−.05	.04	.16	−.01	−.20	−.12	.23	−.09	.07	.14	.12	–
18. Hir. Rec.	59.73	19.15	−.25*	−.13	.09	.40**	.05	−.12	.09	.12	.12	.18	.16	.17	.19	.13	−.02	.14	−.32*

$N = 62$. Gender codes Women = 1, Men = 2

Int. exp. interview experience, *E* extraversion, *O* openness, *N* neuroticism, *A* agreeableness, *C* conscientiousness, *Comm.* communication skills, *Persuas.* persuasion skills, *GMA* general mental ability, *Resp.* responsiveness, *Story* proportion of questions answered with a story, *Pseudo.* proportion of questions answered with a pseudostory, *Exempl.* proportion of questions featuring exemplification, *V/O* proportion of questions answered by expressing a value or an opinion, *Self-d.* proportion of questions answered with a self-description, *Hir. Rec.* hiring recommendation

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 2 Mean proportions (standard deviations) and effect sizes of stories produced in response to behavioral questions, by competency and gender

	Men *M* (*SD*)	Women *M* (*SD*)	Cohen's *d*
Communication	.38 (.50)	.22 (.42)	.32
Persuasion	.38 (.50)	.33 (.48)	.10
Organization	.13 (.34)	.07 (.25)	.18
Stress management	.44 (.51)	.18 (.39)	.51

$N = 62$. Cohen's d is computed using standard deviations for men

between applicant characteristics and the proportion of stories produced. GMA was positively correlated with the proportion of pseudostories and the proportion of exemplifications and negatively correlated with proportion of values/opinions (but visual inspection of the scatterplot revealed this last correlation was due to an outlier on proportion of values/opinions). Moreover, participants higher in conscientiousness and older participants were significantly more likely to produce values/opinions, but the latter correlation was due to an outlier (the same individual as before who was also an outlier in terms of age). To test possible curvilinear relationships, we further computed correlations between the squared measures of applicant characteristics (the Big Five, GMA) and each response type. GMA squared correlated significantly with proportion of stories ($r = -.26$, $n = 62$, $p = .042$), with proportion of pseudostories ($r = .36$, $n = 62$, $p = .004$), and with proportion of exemplifications ($r = .34$, $n = 62$, $p = .006$). Conscientiousness squared correlated significantly with proportion of values/opinions ($r = .28$, $n = 62$, $p = .025$). There were no other correlations between the squared individual-difference variables and response type.

As a final exploration of Research Question 3, we investigated whether the proportion of stories varied by the type of competency and by gender (descriptive statistics in Table 2). We ran a 4 × 2 repeated-measures ANOVA with competency as a four-level within-subjects factor and applicant gender as a two-level between-subjects factor. For within-subjects factors, there was a departure from sphericity, so we report statistics corrected by the Huynh–Feldt method (Field 2009). There was a main effect of competency, $F(2.8, 162.9) = 3.77$, $p = .014$, $\eta_p^2 = .06$. Simple contrasts (with communication as the reference point) indicated that the proportion of stories was lower only for organization. There was also a main effect of gender, $F(1, 59) = 4.1$, $p = .047$, $\eta_p^2 = .07$. Men's proportions of stories were higher than women's ($M = .31$, $SD = .21$ vs. $M = .20$, $SD = .22$). The interaction was not significant, $F(2.8, 162.9) = .72$, *ns*, $\eta_p^2 = .01$.

To answer Research Question 4, we used hierarchical multiple regression to predict hiring recommendations from response types while controlling for all ancillary information available to recruiters. Hiring recommendations were significantly negatively correlated with gender, positively correlated with extraversion, and negatively correlated with the propensity to produce self-descriptions. We computed four regression models. In Model 1, we entered applicant gender as predictor. In Model 2, we additionally entered applicants' scores for the Big Five, communication and persuasion skills, and GMA. In Model 3, we entered applicant responsiveness, because it is correlated with some response types (stories and pseudostories increase responsiveness, see Table 1). In Model 4, we additionally entered applicants' scores for each response type (because the proportions are based on counts and because the resulting distributions were skewed, we applied an arcsine transformation to the raw scores). Results appear in Table 3. The final model shows that men obtained significantly higher hiring recommendations than women. Moreover, producing more stories and pseudostories significantly increases hiring recommendations, while producing more self-descriptions significantly decreases hiring recommendations.

Discussion

This study yields some answers to the research questions we posed. The answer to Research Question 1 (*How do applicants respond to past behavior questions, in particular, how frequent are stories?*) is that applicants rarely respond to past behavior questions by spontaneously producing a story, in fact less than one time in four. Much more frequent are pseudostories, where applicants describe generic events in a summarized way. Our results thus differ markedly from those of Stevens and Kristof (1995), where a large majority of applicants produced stories, many of them even in interviews that did not follow the behavior-description format. However, it is difficult to compare our results with theirs, because of cultural differences in self-presentation between Switzerland and the US (König et al. 2011), as well as the large time gap between our study and theirs, not to mention potential differences in how stories were defined.

The answer to Research Question 2 (*When applicants produce stories, what narrative elements do they contain?*) is that storytelling tends to feature more situational narrative elements than either the actions of the applicants or the results of those actions. Thus, even when they do produce stories, applicants tend to focus more on the context of their actions than on the actions themselves, possibly in order to create some kind of common ground (Clark 1996)

Table 3 Summary of hierarchal multiple regression predicting hiring recommendation

Variable	Model 1		Model 2		Model 3		Model 4	
	B	SE *B*	*B*	SE *B*	*B*	SE *B*	*B*	SE *B*
Intercept	73.37***	7.18	−25.82	44.13	−25.31	42.24	−23.92	39.50
Gender	−10.71*	5.32	−8.37	5.87	−14.17*	6.12	−13.43*	5.76
E			14.30*	5.86	13.32*	5.62	7.77	5.36
A			−2.57	5.87	−4.39	5.66	−5.77	5.39
O			1.25	4.52	−1.50	4.47	−1.78	4.18
N			−.01	5.02	2.11	4.88	2.74	4.53
C			3.24	4.69	1.42	4.55	2.55	4.25
Comm.			−3.18	5.79	−.28	5.69	2.31	5.29
Persuas.			5.97	5.02	7.01	4.83	3.74	4.63
GMA			29.47	18.11	31.54	17.36	22.07	18.05
Resp.					.113*	.047	.076	.055
Story							28.27**	9.56
Pseudo.							22.52*	8.70
Exempl.							−1.29	7.10
V/O							5.74	5.97
Self-d.							−10.83*	5.03
Adj. R^2		.05		.13		.20		.33
F for R^2 change		4.05*		1.70		5.76*		2.93*

$N = 62$. Gender: 1 = women, 2 = men

E extraversion, *A* agreeableness, *O* openness, *N* neuroticism, *C* conscientiousness, *Comm.* communication skills, *Persuas.* persuasion skills, *GMA* general mental ability, *Resp.* responsiveness, *Story* proportion of questions answered with a story, *Pseudo.* proportion of questions answered with a pseudostory, *Exempl.* proportion of questions featuring exemplification, *V/O* proportion of questions answered by expressing a value or an opinion, *Self-d.* proportion of questions answered with a self-description, *Hir. Rec.* hiring recommendation

* $p < .05$; ** $p < .01$; *** $p < .001$

with the recruiter. The focus on narrating situational elements may be a pragmatic necessity in telling a story to another person whom one does not know; however, the same logic would seem to hold for the applicants' actions and their effects.

The answer to Research Question 3 (*Is story production related to applicants' personal characteristics?*) is that there were some relations between the characteristics we examined (personality, self-evaluated persuasion or communication skill, GMA, age, gender, interview experience) and the proportion of stories produced. We found an effect of gender on the propensity to produce stories. Moreover, the effect size of gender differences varied by competency, with a medium effect size for stress management and a small effect size for persuasion or organization. It is unclear to what this is due, and it is unclear whether it is a gender effect *per se* or due to the situation (the interviewer was female).

We also found a significant correlation between conscientiousness and the tendency to voice values/opinions (both linear and curvilinear relationships). This may be because people high in conscientiousness are achievement-oriented and have a stronger sense of duty. We also found a negative curvilinear relation between GMA the propensity to produce stories, and positive (linear and curvilinear) relations between GMA and the propensity to tell pseudostories and use exemplification. These results suggest that responding to past behavior questions may involve a cognitive component. While both stories and pseudostories have a narrative component (both being focused on events), pseudostories are more abstract and decontextualized than stories. The propensity of higher-GMA applicants to produce more pseudostories and fewer stories may reflect a spontaneous tendency to abstract construal of experience.

We also found that the type of competency affects storytelling: questions about organizing competencies elicited fewer stories than those about the other competencies. The way this competency gets used in everyday work life may be less conducive to storytelling, because organizing one's work is by nature a more routine activity for which applicants may not have access to a particular episode or event where they demonstrated this competency. More generally, this finding raises the question of potential baseline differences in how easy it is to tell a story about different competencies. In other words, just as some personality traits are easier to observe than others (Funder

1995), some competencies may be easier to translate into a story than others.

The answer to Research Question 4 (*How are applicants' responses related to interview outcomes?*) is that three response types, i.e., stories, pseudostories, and self-descriptions, predicted hiring recommendations of professional recruiters. This is a strong finding, because it emerges despite having controlled for a range of ancillary information recruiters had about applicants (e.g., gender, self-report data, GMA). Recruiters gave higher recommendations for applicants who produced more stories and pseudostories. This suggests that recruiters are sensitive to the presence of narrative content in applicants' responses to past behavior questions. Moreover, self-descriptions, which are decontextualized statements, decreased hiring recommendations. Although we do not know how these response types affect hiring recommendations (i.e., what processes mediate the link), the fact that narrative content increases recommendations and non-narrative content decreases them suggests that narrative content is an expected dimension of applicants' responses to past behavior questions. Moreover, self-descriptions (*I'm a good communicator*) may seem less credible to recruiters if they are not backed up by concrete evidence.

This study has some limitations. One limitation is the sample of young, inexperienced applicants. Applicants who are older or who have more work experience or interviewing experience may evidence different patterns of responses to past behavior questions. Thus, the dataset was not ideal for testing the effects of some individual differences. Another limitation is the nature of the position. Although applicants knew they were applying for a real job, they may not have been motivated to excel as much as for a longer-term engagement or a professional position. A third limitation is the relatively small sample size, which limits statistical power to detect some weak relations, as well as the exploratory nature of our study. On the other hand, a sample of 62 interviews conducted in a fully standardized manner for the same position is rare to find. The standardized interview situation where the interviewer systematically refrained from any kind of prompt constitutes a strength insofar as it creates a questioning context free from confounds related to the interviewer's behavior, thus enabling measurement of a "baseline" of storytelling propensity. However, such a standardized interview situation is hardly representative of the diversity of interviews in practice and also constitutes a limitation. As a result of these limitations, more research is needed before generalizing our findings, especially as other data sets (Stevens and Kristof 1995) have found a higher incidence of storytelling.

This study has several implications for research and practice. Perhaps the most important implication is the number of questions it opens up for research on social interaction in the selection interview. More research is needed in order to develop scientific knowledge about how applicants respond to selection interview questions. Research might focus on identifying conditions that affect the production of stories: applicant individual differences, question formulation, recruiter behavior, and type of competency seem to be important factors. Research might also focus on the content of stories and the question of story quality, i.e., what constitutes a good story in a selection context, both from the point of view of the applicant and the recruiter (Ralston et al. 2003). From a practical point of view, there are implications for training for both interviewers and applicants. Past behavior questions do not always successfully elicit detailed stories. Recruiters using structured interviews should be made aware of this fact. They might also be trained in techniques to elicit stories, or to help applicants transform pseudostories or abstract self-descriptions into stories (e.g., via follow-up questioning). Training programs for applicants could also focus on helping them think about their experiences in ways that can be framed as good stories (Ralston et al. 2003); i.e., stories that express in an accurate and detailed manner their level of mastery of a specific competence.

Acknowledgments Research was funded by the Swiss National Science Foundation (Sinergia project *Interactional Competences in Institutional Practices: Young People between School and the Workplace,* CRSII1_136291). The data were obtained from another Swiss National Science Foundation Sinergia project, SONVB, FNCRSII2-127542/1. We thank Dr. Daniel Gatica-Perez, IDIAP; Dr. Marianne Schmid Mast, University of Neuchatel; and Dr. Tanzeem Choudhury, Cornell University, for granting us access to their data. We thank Franciska Krings and Franziska Tschan for advice on an earlier version of the manuscript.

References

Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science, 27*, 195–225.

Bangerter, A., Clark, H. H., & Katz, A. R. (2004). Navigating joint projects in telephone conversations. *Discourse Processes, 37*, 1–23.

Bangerter, A., Mayor, E., & Pekarek Doehler, S. (2011). Reported speech in conversational narratives during nursing shift handover meetings. *Discourse Processes, 48*, 183–213.

Bangerter, A., Roulin, N., & König, C. J. (2012). Personnel selection as a signaling game. *Journal of Applied Psychology, 97*, 719–738.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology, 79*, 941–952.

Boje, D. M. (2008). *Storytelling organizations*. London: Sage.

Bruner, J. (1990). *Acts of meaning*. Cambridge: Harvard University Press.

Button, G., & Casey, N. (1984). Generating topic: the use of topic initial elicitors. In M. Atkinson & J. Heritage (Eds.), *Structures*

of social action (pp. 167–190). Cambridge: Cambridge University Press.

Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50*, 655–702.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Schober, M. F. (1991). Asking questions and influencing answers. In J. M. Tanur (Ed.), *Questions about questions: Inquiries into the cognitive bases of surveys* (pp. 15–48). New York: Russell Sage Foundation.

Conrad, F. G., & Schober, M. F. (2005). Promoting uniform question understanding in today's and tomorrow's surveys. *Journal of Official Statistics, 21*, 215–231.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*, 5–13.

Denning, S. (2004). Telling tales. *Harvard Business Review, 82*, 122–129.

Edwards, D., & Potter, J. (1993). Language and causation: A discursive action model of description and attribution. *Psychological Review, 100*, 23–41.

Ellis, A. P. J., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type. *Journal of Applied Psychology, 87*, 1200–1208.

Field, A. (2009). *Discovering statistics using SPSS*. Newbury Park, CA: Sage.

Frauendorfer, D., Schmid Mast, M., Nguyen, L., & Gatica-Perez, D. (2013a). Predicting hiring decision via automatic social sensing of job applicant nonverbal interview behavior (Manuscript submitted for publication).

Frauendorfer, D., Schmid Mast, M., Nguyen, L., & Gatica-Perez, D. (2013b). A step towards automatic applicant selection: Predicting job performance based on applicant nonverbal interview behavior (Manuscript submitted for publication).

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652–670.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225–246). Cambridge: Cambridge University Press.

Goodwin, C. (1987). Forgetfulness as an interactive resource. *Social Psychology Quarterly, 50*, 115–130.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (pp. 41–58). New York: Academic Press.

Habermas, T., & Bluck, S. (2000). Getting a life: The emergence of the life story in adolescence. *Psychological Bulletin, 126*, 748–769.

Holt, E. (1996). Reporting on talk: The use of direct reported speech in conversation. *Research on Language and Social Interaction, 29*, 219–245.

Huffcutt, A. I., & Arthur, W., Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184–190.

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., et al. (2005). A conditional reasoning measure for aggression. *Organizational Research Methods, 8*, 69–99.

Jetter, W. (2008). *Effiziente Personalauswahl: Die richtigen Mitarbeiter mit strukturierten Einstellungsinterviews finden*. Stuttgart: Schäffer-Poeschel.

Kemper, S. (1984). The development of narrative skills: Explanations and entertainments. In S. A. Kuczaj (Ed.), *Discourse development* (pp. 99–124). New York: Academic Press.

Kessler, R. (2006). *Competency-based interviews*. Pompton Plains, NJ: Career Press.

König, C. J., Hafsteinsson, L. G., Jansen, A., & Stadelmann, E. H. (2011). Applicants' self-presentational behavior across cultures: Less self-presentation in Switzerland and Iceland than in the US. *International Journal of Selection and Assessment, 19*, 331–339.

Labov, W., & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helms (Ed.), *Essays on the verbal and visual arts* (pp. 12–44). Seattle, WA: University of Washington Press.

Larson, W. W. (2001). *Ten-minute guide to conducting a job interview*. Indianapolis, IN: Macmillan.

Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment, 14*, 299–316.

Lievens, F., & De Paepe, A. (2004). An empirical investigation of interviewer-related factors that discourage the use of high structure interviews. *Journal of Organizational Behavior, 25*, 29–46.

Maurer, T., Solamon, J., & Lippstreu, M. (2008). How does coaching interviewees affect the validity of a structured interview? *Journal of Organizational Behavior, 29*, 355–371.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79*, 599–616.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology, 60*, 683–729.

Norrick, N. (2000). *Conversational narrative: Storytelling in everyday talk*. Amsterdam: John Benjamins.

Orr, J. (1996). *Talking about machines: An ethnography of a modern job*. Ithaca, NY: Cornell University Press.

Posthuma, R., Morgeson, F., & Campion, M. (2002). Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time. *Personnel Psychology, 55*, 1–81.

Prabakar Kamath, S. (2009). *Competency-based interviewing*. New Delhi: Excel Books.

Pratt, M. W., & Robins, S. L. (1991). That's the way it was: Age difference in the structure and quality of adults' personal narratives. *Discourse Processes, 14*, 73–85.

Ralston, S. M., Kirkwood, W. G., & Burant, P. A. (2003). Helping interviewees tell their stories. *Business Communication Quarterly, 66*, 8–22.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51*, 649–660.

Roulin, N., & Bangerter, A. (2012). Understanding the academic-practitioner gap for structured interviews: "Behavioral" interviews diffuse, "structured" interviews do not. *International Journal of Selection and Assessment, 20*, 149–158.

Roulin, N., Bangerter, A., & Wüthrich, U. (2012). *Réussir l'entretien d'embauche comportemental: La méthode pour identifier et sélectionner les futurs employés performants. [Succeeding at the behavioral interview: The method for identifying and selecting future high performers.]*. Bruxelles: De Boeck Professionals.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289–327.

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations*. Monterey, CA: Brooks/Cole.

Schmid Mast, M. (2002). Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research, 28*, 420–450.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Schuler, H. (1992). Das multimodale einstellungsinterview. *Diagnostica*, *38*(4), 281–300.

Sidnell, J. (2006). Coordinating gesture, talk, and gaze in reenactments. *Research on Language and Social Interaction, 39*, 377–409.

Silvester, J. (1997). Spoken attributions and candidate success in graduate recruitment interviews. *Journal of Occupational and Organizational Psychology, 70*, 61–73.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Blackwell.

Stevens, C. K., & Kristof, A. L. (1995). Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology, 80*, 587–606.

Suchman, L., & Jordan, L. (1990). Interactional troubles in face-to-face survey interviews. *Journal of the American Statistical Association, 85*, 232–253.

Van der Zee, K. I., Bakker, A. B., & Bakker, P. (2002). Why are structured interviews so rarely used in personnel selection? *Journal of Applied Psychology, 87*, 176–184.

Wonderlic, E. G. (Ed.). (2001). *Wonderlic personnel test manual*. Northfield, IL: E.F. Wonderlic & Assoc.